Molecular Evolution and Mosaicism of Leptospiral Outer Membrane Proteins Involves Horizontal DNA Transfer

David A. Haake,[1,2]* Marc A. Suchard,[3] Melissa M. Kelley,[1,2] Manjula Dundoo,[1,2]
David P. Alt,[4] and Richard L. Zuerner[4]

*Division of Infectious Diseases, Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, California 90073[1];
Department of Medicine,[2] and Department of Biomathematics,[3] David Geffen School of Medicine at the University
of California—Los Angeles, Los Angeles, California 90095; and National Animal Disease Center,
Agricultural Research Service, U.S. Department of Agriculture, Ames, Iowa 50010[4]*

Leptospires belong to a genus of parasitic bacterial spirochetes that have adapted to a broad range of mammalian hosts. Mechanisms of leptospiral molecular evolution were explored by sequence analysis of four genes shared by 38 strains belonging to the core group of pathogenic *Leptospira* species: *L. interrogans, L. kirschneri, L. noguchii, L. borgpetersenii, L. santarosai,* and *L. weilii*. The 16S rRNA and *lipL32* genes were highly conserved, and the *lipL41* and *ompL1* genes were significantly more variable. Synonymous substitutions are distributed throughout the *ompL1* gene, whereas nonsynonymous substitutions are clustered in four variable regions encoding surface loops. While phylogenetic trees for the 16S, *lipL32,* and *lipL41* genes were relatively stable, 8 of 38 (20%) *ompL1* sequences had mosaic compositions consistent with horizontal transfer of DNA between related bacterial species. A novel Bayesian multiple change point model was used to identify the most likely sites of recombination and to determine the phylogenetic relatedness of the segments of the mosaic *ompL1* genes. Segments of the mosaic *ompL1* genes encoding two of the surface-exposed loops were likely acquired by horizontal transfer from a peregrine allele of unknown ancestry. Identification of the most likely sites of recombination with the Bayesian multiple change point model, an approach which has not previously been applied to prokaryotic gene sequence analysis, serves as a model for future studies of recombination in molecular evolution of genes.

The genus *Leptospira* consists of a diverse group of pathogenic and saprophytic spirochetes, currently classified into 17 genomospecies based on DNA-DNA hybridization studies (25). Species assignments are consistent with phylogenetic analysis based on 16S gene sequences (19, 36). A universal feature of pathogenic leptospires is the ability to parasitize the proximal renal tubules of a wide variety of wild and domestic animals. Infection with host-adapted leptospiral serovars can result in lifelong renal carriage and urinary shedding. In humans, exposure to infected host animals or contaminated water or soil results in potentially lethal disease. Leptospirosis is now recognized as the most widespread zoonosis (25). Leptospiral diversity is reflected by the broad array of mammals that can serve as reservoir hosts and by the fact that over 200 leptospiral serovars have been described (8). Changes in the antigenic composition of lipopolysaccharide are thought to account for serovar diversity (50). However, the molecular diversity of surface-exposed leptospiral outer membrane proteins (OMPs) has not previously been examined.

Several leptospiral OMPs have been described, differing in their degrees of surface exposure. OmpL1 is a porin that has been shown to be surface exposed by immunoelectron microscopy and surface immunoprecipitation (39). Like porins of gram-negative bacteria, OmpL1 is an oligomer and its electrophoretic migration in acrylamide gels is heat modifiable (39).

Based upon these findings and beta-moment analysis of the OmpL1 amino acid sequence, a topological model of OmpL1 was proposed with five surface-exposed loops and 10 beta-sheet transmembrane segments (13). In addition to OmpL1, a number of lipoprotein OMPs have been characterized, including LipL21 (6), LipL32 (14), LipL36 (16), and LipL41 (40), designated according to their apparent molecular weights determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. LipL21 and LipL41 appear to be surface exposed, whereas LipL36 is thought to be restricted to the inner leaflet of the outer membrane (6, 40). OmpL1 and LipL41 exhibit synergistic immunoprotection in the hamster model of leptospirosis (17). LipL32 is the leptospiral major OMP and has an immunoprotective effect when hamsters are immunized with LipL32 incorporated into an adenovirus construct (4). The degree to which LipL32 is surface exposed has not been determined. LipL32, LipL41, and OmpL1 are major antigens in the humoral immune response to leptospirosis (10, 12). Recently, a new family of leptospiral lipoproteins was described which contain repeated bacterial immunoglobulin-like domains (28). Expression of two of these Lig (leptospiral immunoglobulin-like) proteins, LigA and LigB, is associated with virulence. The *ligC* gene, encoding a third member of the family, appears to be a pseudogene in *Leptospira interrogans* and *Leptospira kirschneri*. A third class of leptospiral OMP are the LipL45-derived peripheral membrane proteins associated with, but not integrated into, the outer membrane (29, 31).

Comparative analysis of OMP gene sequences can reveal insights into novel mechanisms of molecular evolution in pathogenic bacteria. Previous studies have revealed that the

---

* Corresponding author. Mailing address: Division of Infectious Diseases, 111F, 11031 Wilshire Blvd., VA Greater LA Healthcare System, Los Angeles, CA 90073. Phone: (310) 268-3814. Fax: (310) 268-4928. E-mail: dhaake@ucla.edu.

TABLE 1. Strains of *Leptospira* species used in DNA sequence analysis studies[a]

Species	Serogroup	Serovar	Strain[b]	Isolation country	Source
L. borgpetersenii	Autumnalis	Srebarna	1409/69 (AT15)	Bulgaria	Shrew
	Ballum	Ballum	Mus 127 (BM01)	Denmark	Field mouse
	Hebdomadis	Nona	Nona (HB02)	Zaire	Human
	Javanica	Ceylonica	Piyasena (JV02)	Sri Lanka	Human
	Javanica	Javanica	Veldrat Batavia 46 (VB46)	Indonesia	Field rat
	Mini	Mini	Sari (HB10)	Italy	Human
	Sejroe	Hardjo	93U (H15B)	—[c]	—
	Tarassovi	Tarassovi	Perepelicin (TA01)	Russia	Human
L. interrogans	Australis	Bratislava	Jez-Bratislava (AS05)	Czechoslovakia	Hedgehog
	Autumnalis	Autumnalis	Akiyami (AT01)	Japan	Human
	Bataviae	Bataviae	Van Tienen (BT02)	Indonesia	Human
	Canicola	Canicola	Hond Utrecht (CA01)	The Netherlands	Dog
	Icterohaemorrhagiae	Copenhageni	L1-130 (L130)	Brazil	Human
	Icterohaemorrhagiae	Icterohaemorrhagiae	RGA (RGA)	Belgium	Human
	Hebdomadis	Kremastos	Kremastos (HB04)	Australia	Human
	Icterohaemorrhagiae	Lai	56601 (LLai)	China	Sewage
	Pomona	Pomona	Pomona (PO01)	Australia	Human
	Sejroe	Hardjo	Hardjoprajitno (H15A)	Indonesia	Human
L. kirschneri	Australis	Ramisi	Musa (AS11)	Kenya	Human
	Autumnalis	Erinaceiauriti	Erinaceus Auritus 670 (AT04)	Russia	Hedgehog
	Bataviae	Djatzi	HS26 (BT04)	Puerto Rico	Human
	Canicola	Galtoni	LT1014 (CA02)	Argentina	Cow
	Hebdomadis	Kambale	Kambale (HB03)	Zaire	Human
	Grippotyphosa	Grippotyphosa	RM52 (GR01)	United States	Pig
	Pomona	Mozdok	5621 (PO04)	Russia	Field vole
L. noguchii	Australis	Nicaragua	1011 (AS10)	Nicaragua	Least weasel
	Louisiana	Orleans	LSU2580 (AT10)	United States	Nutria
	Pomona	Proechimys	LT796 (PO02)	—	—
L. santarosai	Bataviae	Rioja	MR12 (BT10)	Peru	Opossum
	Grippotyphosa	Canalzonae	CZ288 (GR05)	—	—
	Mini	Georgia	LT117 (HB13)	United States	Raccoon
	Pomona	Tropica	CZ299 (PO05)	Panama	Spiny rat
	Pyrogenes	Alexi	HS-616 (PY07)	Puerto Rico	Human
	Sejroe	Trinidad	LT1098 (HB20)	—	—
	Shermani	Shermani	LT821 (SH01)	—	—
	Tarassovi	Bakeri	LT79 (TA02)	United States	Opossum
L. weilii	Javanica	Coxi	Cox (JV05)	Malaysia	Human
	Tarassovi	Vughia	LT89-68 (TA15)	Vietnam	Human

[a] Thirty-eight strains were included in the analysis of the 16S, *lipL32, lipL41,* and *ompL1* genes.
[b] The four-character strain code is in parentheses.
[c] —, information not available.

organization of leptospiral genomes is relatively fluid due to various types of recombination events. Comparative pulsed-field gel electrophoresis studies of even closely related strains reveal many large rearrangements (3, 52). In addition, *Leptospira* species contain repetitive transposase-encoding insertion sequences (IS), some of which may play a role in producing genomic rearrangements (3, 24). The mobility of one of these IS elements, IS*1500*, is sufficient to allow Southern blot discrimination of 15 different groups within isolates of *L. interrogans* serovar Pomona type kennewicki (51). Evidence for horizontal transfer of DNA among *Leptospira* species comes from studies of the intervening sequences found within the 23S rRNA gene (37) and from the finding that the leptospiral lipopolysaccharide biosynthetic locus (*rfb*) is located in a genomic island that was probably acquired through horizontal transfer from a gram-negative source (24). Comparative sequence analysis of *rfb* loci has provided information about

variability of genes involved in lipopolysaccharide biosynthesis (7). However, relatively little is known about the mechanisms and extent of molecular diversity of leptospiral OMP genes. Given the evidence that OmpL1, LipL41, and LipL32 play a role in protective immunity, the degree of amino acid sequence variation of these OMPs was determined in a large number of strains. This information is useful in making predictions about how broadly protective these immunogens would be and in understanding the frequency of recombination events involving OMP genes.

(Portions of this work were presented at the 102nd Annual Meeting of the American Society for Microbiology, Washington, D.C., 21 to 25 May 2002.)

MATERIALS AND METHODS

Bacterial strains and media. Leptospiral strains (Table 1) were obtained from the National Leptospirosis Reference Center (National Animal Disease Center,

TABLE 2. PCR and sequencing primers[a]

Primer	Sequence (5′–3′)	Position[b] (nucleotides)
16S	GAGAGTTTGATYCTGGCTCAG	7–27
16S	ACGGGCGGTGTGTACAAGG	1353–1371 c′
16S	ATCCTCATGGCCTTTATGTC	1163–1182
16S	AGTGAACGGGATTAGATACC	747–766
16S	ACACGGTCCATACTCCTACG	312–331
16S	CCTAGACATAAAGGCCATGA	1167–1186 c′
16S	CCCGTCAATTCCTTTGAGTT	877–896 c′
16S	CTGCTGCCTCCCGTAGGAGT	323–342 c′
LipL32F∗	ATGAAAAAACTTTCGATTTTG	1–21
LipL32R∗	TTACTTAGTCGCGTCAGAAGC	799–819 c′
LipL32F1	GAGTTCGTATGATTTCCCCAA	280–301
LipL32R1	GTATCAAACCAATGTGGCATT	378–398 c′
LipL32F2∗	CTAAGTTCATACCGTGATTT	−39−−19
LipL32R2∗	ATTACTTAGTCGCGTCAGAA	801–819 c′
LipL32F3	TCCTCCAAAATCTTTTGAC	564–582
LipL32R3	GGAATCCAAACATAGAGATA	241–260 c′
LipL41F	TGTTACCCATGGGGAGAAAATTATCTTCTC	3–20
LipL41R	AAAGGACTCGAGTTACTTTGCCGTTGCCTTTC	1052–1069 c′
LipL41F1	CGCTTCTATGGCAACTGGTAA	478–498
LipL41R1	TTAATGAGAGTAGCATCGAGA	562–582 c′
LipL41F2∗	ATGAGAAAATTATCTTCTCTA	1–22
LipL41R2∗	TTACTTTGCGTTGCTTTCGTC	1049–1069 c′
OmpL1F	ATGATCCGTAACATAAGTAAGGCATTG	1–27
OmpL1R	TTAGAGTTCGTGTTTATAACCGAATCT	937–963 c′
OmpL1F1	CCTCGATAGAACTACTGGCGG	312–332
OmpL1R1	AGGAGCAATTCCAGAAGTTCT	745–765 c′
OmpL1F2∗	ATGATCCGTAACATAAGTAAG	1–17
OmpL1R2∗	TTAGGTTCGTGTTTATAACC	943–963 c′

[a] PCR primers are indicated by asterisks.
[b] c′, complementary strand.

Agricultural Research Service, U.S. Department of Agriculture, Ames, Iowa). Most of the leptospiral strains used in this study are also described in a recent DNA relatedness study (5). Leptospires were cultivated in Johnson-Harris bovine serum albumin Tween 80 medium (23).

DNA sequence determination. The 16S rRNA, *lipL32*, *lipL41*, and *ompL1* gene sequences were obtained by direct sequencing of PCR amplicons generated with *Tth* polymerase (Clonetech, Palo Alto, Calif.). Some of the genes analyzed have been previously reported (13, 14, 19, 38, 40). Previously unpublished sequences have been submitted to GenBank (see below). Primers used for sequencing are listed in Table 2.

DNA sequence analysis. The 16S rRNA, *lipL32*, *lipL41*, and *ompL1* genes from a total of 38 strains were examined. Codon adaptation indices were calculated by DNA Strider. Sequence alignments were performed with Clustal W (46) with its default settings. Alignments of the 16S rRNA data were based on nucleotide sequences, and alignments of the genes *lipL32*, *lipL41*, and *ompL1* were based on amino acid translations to maintain reading frames. Phylogenetic analyses were performed by using a probabilistic approach in a Bayesian framework, as this approach enabled us to employ a rich class of statistical models to explore the possibility of intragenic recombination and to identify amino acid sites driven by positive selection. Accessible reviews are available for readers unfamiliar with Bayesian phylogenetic and recombinational analysis (18, 21, 41). Phylogenetic reconstructions of 16S rRNA, *lipL32*, and *lipL41* and initial analyses of *ompL1* were performed by using Markov chain Monte Carlo computation in the program MrBayes (20) and included multiple independent simulations to assess adequate estimation of the phylogenetic tree (45).

Possible intragenic recombination in *ompL1* was examined by using a novel extension to the Bayesian multiple change point (BMCP) model proposed by Suchard et al. (43, 44). This model assumes that the sites along a sequence alignment separate into an unknown number of contiguous segments, each with possibly different evolutionary relationships between the organisms, evolutionary rates, and transition-transversion ratios. Differing evolutionary relationships on

either side of a breakpoint suggests recombination (26). One strength of the BMCP model is that it can measure uncertainty in breakpoint locations, determine the most likely parental sequences of the putative recombinant, and assess the statistical significance of recombination simultaneously; this simultaneous approach avoids a statistical pitfall inherent in sequential testing for recombination found in many recombination detection programs (44).

Current implementations of the BMCP model limit analysis to five sequences (33) or assume a known and fixed phylogenetic tree relating a nearly unlimited number of representative parental sequences (44). In this study, these assumptions were relaxed to allow for an a priori unknown relationship between any number of sequences (n). Among n sequences, there exists $E = (2n − 5)(2n − 3). . . (1)$ possible trees. The BMCP model was adapted in two steps. In step one, the original (L) alignment sites for a gene from the n sequences were chopped up into a large number (M) of overlapping contiguous sections, each of length l. The MrBayes program was then employed to estimate the 5 most probable trees for each section, and a list of these trees across all sections was created. Assume this process resulted in a list of P unique trees. In step two, the tree at each site was allowed to equal any one of the P possible trees. Although the total number of allowable tree configurations (P^L) is, in general, considerably smaller than E^L, the number of total possible configurations in an unrestricted state-space, little bias is expected. Assuming the initial list of P trees is overly diffuse compared to the true posterior distribution over all possible trees in each inferred segment, results from the extended BMCP should well approximate the true posterior distribution.

This extended BMCP model was employed to identify a collection of *ompL1* sequences with significant support against recombination and in favor of a single tree describing their evolutionary relationships along the entire gene. Sequences were included in this collection from as many genomospecies as possible. These nonrecombinant sequences with their backbone tree then served as possible parental representatives for the analyses of the putative *ompL1* recombinant sequences. For each putative recombinant, breakpoints and parental represen-

tatives were inferred and their support in favor of recombination was assessed by using the approach of Suchard et al. (43).

Amino acid sites in *ompL1* subject to positive selection were identified by codon substitution models (11) for estimating synonymous and nonsynonymous rates in PAML (48, 49). In particular, models were employed in which the ratio of nonsynonymous to synonymous changes (ω) may vary from codon to codon. Statistical significance of positively selected sites was assessed by an empirical Bayes approach under model M8 (48). To control for possible intragenic recombination when performing this analysis, the gene alignment was first partitioned into independent segments based on the most probable breakpoint locations. For each segment, the tree was set equal to that most probable for the segment and tests for positive selection proceeded assuming independence between segments. Results are compared to those obtained without controlling for recombination.

Nucleotide sequence accession number. Previously unpublished gene sequences were submitted to GenBank and assigned accession numbers AY461855 to AY462006.

RESULTS

DNA sequence conservation rates. The genes encoding 16S, *lipL32*, *lipL41*, and *ompL1* were sequenced from each of 38 strains (Table 1), yielding alignments of 1,186, 808, 933, and 795 bp, respectively. The 16S, *lipL32*, and *lipL41* genes are highly conserved, with 99.0, 96.9, and 90.9% average pairwise DNA sequence identity, respectively. Most of the *lipL32* and *lipL41* single-nucleotide polymorphisms were synonymous. In contrast, the *ompL1* gene is more variable, with 85.1% average pairwise DNA sequence identity (Fig. 1), and a majority of *ompL1* single-nucleotide polymorphisms were nonsynonymous. Amino acid sequence variations appear evenly distributed in LipL32 and LipL41 proteins, whereas those of OmpL1 appear clustered in four variable regions (Fig. 2A). Our sequencing strategy did not allow us to evaluate the first 24 amino acids (the signal peptide) or the last 20 amino acids in OmpL1. Of the 276 evaluable amino acids in OmpL1, 69 (25%) were variable. The topological model of OmpL1 indicates that 86% (59 of 69) of the variable amino acids occur in its surface-exposed loops (Fig. 2B). Significant (two-tailed *t* test for paired samples; *P* < 0.0001) differences were found in the DNA sequence variability of the four genes and in the amino acid sequence variability of the three translated genes. The codon adaptation indices for the *ompL1*, *lipL41*, and *lipL32* genes were 0.224, 0.330, and 0.334, respectively, indicating that, of these three genes, *ompL1* exhibits the highest degree of codon bias.

Description of phylogenetic trees. Figure 3A depicts the posterior consensus tree inferred from the 16S gene sequences. The tree describes the evolutionary relatedness of the 38 sequences, and its branch patterns follow the species assignments determined by prior DNA-DNA hybridization studies (5). The six *Leptospira* species partition into two clusters: an *L. borgpetersenii*, *L. santarosai*, and *L. weilii* cluster and an *L. interrogans*, *L. kirschneri*, and *L. noguchii* cluster. Within the former cluster, all of the sequences of each of the three species form monophyletic groups. In contrast, the relatedness tree of the latter cluster exhibits a more complex pattern. The posterior probability that the *L. kirschneri* sequences form a monophyletic group is <0.03. Instead, the *L. interrogans* sequences form a monophyletic group within the *L. kirschneri* group. Similarly, the posterior probability that the *L. noguchii* sequences form a monophyletic group is <0.04. Instead, the *L. interrogans* and *L. kirschneri* sequences form a monophyletic group within the *L. noguchii* branch.
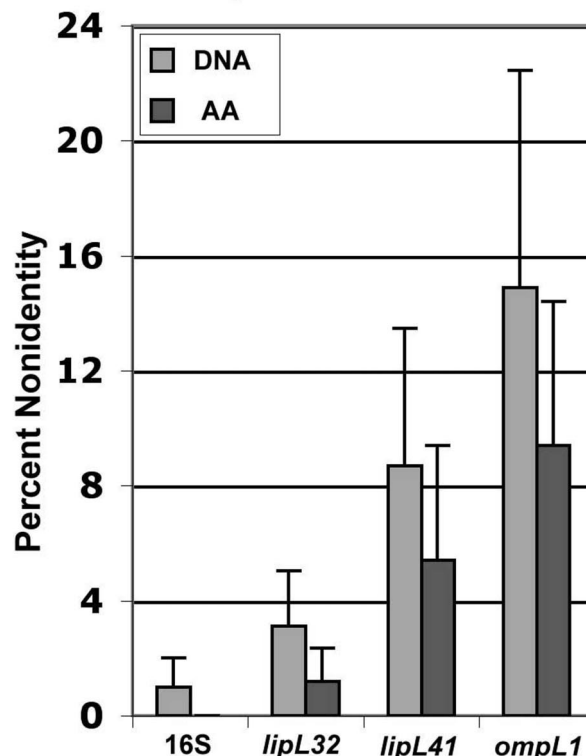

FIG. 1. Comparison of DNA and amino acid (AA) sequence non-identity among leptospiral genes. Sequences of the 16S, *lipL32*, *lipL41*, and *ompL1* genes from 38 leptospiral strains were compared, revealing striking differences in the degree of DNA and amino acid sequence variability between genes. The 16S and *lipL32* genes were less variable, and most *lipL32* and *lipL41* nucleotide mutations were synonymous. In contrast, significantly greater DNA and amino acid sequence variability was observed for the *lipL41* and *ompL1* genes. Differences in DNA and amino acid sequence variability were significant for all genes (two-tailed *t* test for paired samples, *P* < 0.0001). Error bars indicate standard deviations.

The inferred *lipL32* tree (Fig. 3B) generally mirrored the 16S gene tree. As in the 16S tree, the *L. interrogans* group emerges from within the *L. kirschneri* branch. However, in contrast to the 16S tree, the *L. noguchii* sequences form a monophyletic group in the *lipL32* tree (posterior probability, >0.99). This indicates a greater degree of separation of the *L. noguchii* sequences from the *L. interrogans* and *L. kirschneri* sequences. Another contrast with the 16S sequence tree is that the *lipL32* sequences of *L. borgpetersenii* and *L. weilii* do not form separate monophyletic groups (posterior probability, >0.99). This is consistent with the finding that the *lipL32* sequence of *L. weilii* strain JV05 was identical to that of four *L. borgpetersenii* strains.

The inferred *lipL41* tree (Fig. 3C) is distinct from those of the 16S and *lipL32* sequence trees in two regards. First, in the *lipL41* tree, all six species form monophyletic groups, consistent with the greater degree of sequence variation for the
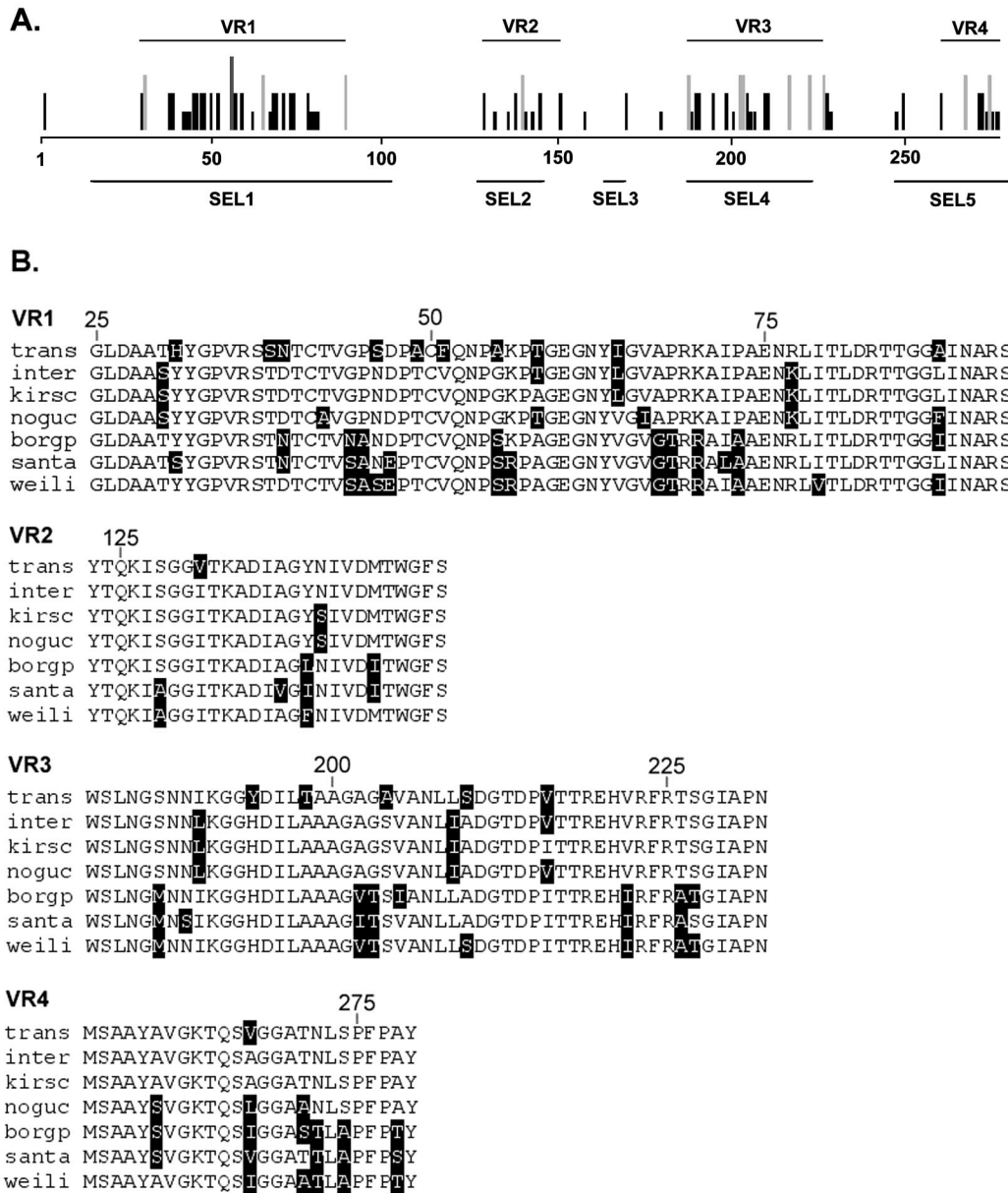
FIG. 2. Localization of OmpL1 variable regions. Comparison of the OmpL1 amino acid sequences of 38 leptospiral strains reveals four variable regions corresponding to the four largest surface-exposed loops. (A) Histogram of OmpL1 amino acid sequence variability. Variable amino acids are clustered in four variable regions corresponding to the four largest surface-exposed loops of OmpL1. Variable region 1 (VR1) is located in the first surface-exposed loop (SEL1) and contains the highest number of variable amino acids. Variable region 2 (VR2) is located in the second surface-exposed loop (SEL2), variable region 3 (VR3) is located in the fourth surface-exposed loop (SEL4), and variable region 4 (VR4) is located in the fifth surface-exposed loop (SEL5). The height of each bar in the histogram indicates the degree of amino acid sequence variation from the consensus amino acid residue at that location. The OmpL1 amino acid sequence is numbered from the N-terminal amino acid of the mature protein. (B) Alignments of OmpL1 variable regions. Alignments of the four OmpL1 variable regions are shown by using the consensus amino acid sequences for the horizontally transferred peregrine allele (trans) and each of the six *Leptospira* species examined in this study: *L. interrogans* (inter), *L. kirschneri* (kirsch), *L. noguchii* (noguc), *L. borgpetersenii* (borgp), *L. santarosai* (santa), and *L. weilii* (weili). Locations of variable amino acids are indicated by dark boxes. The OmpL1 amino acid sequence is numbered from the N-terminal amino acid of the mature protein.

lipL41 sequences. Second, the *lipL41* sequence of *L. borgpetersenii* strain HB10 is no longer grouped with *L. borgpetersenii* (posterior probability, >0.99) and is instead most likely grouped with *L. interrogans* (posterior probability, >0.92). This suggests that horizontal transfer of the strain HB10 *lipL41* sequence may have occurred. The entire *lipL41* sequence of *L.*

borgpetersenii strain HB10 is characteristic of *L. interrogans* *lipL41* sequences, indicating that the entire gene was acquired during the putative transformation event.

The inferred *ompL1* tree (Fig. 3D) also differs strikingly from the 16S, *lipL32*, and *lipL41* trees in that there is formation of a major new branch of the tree involving 20% (8 of 38) of
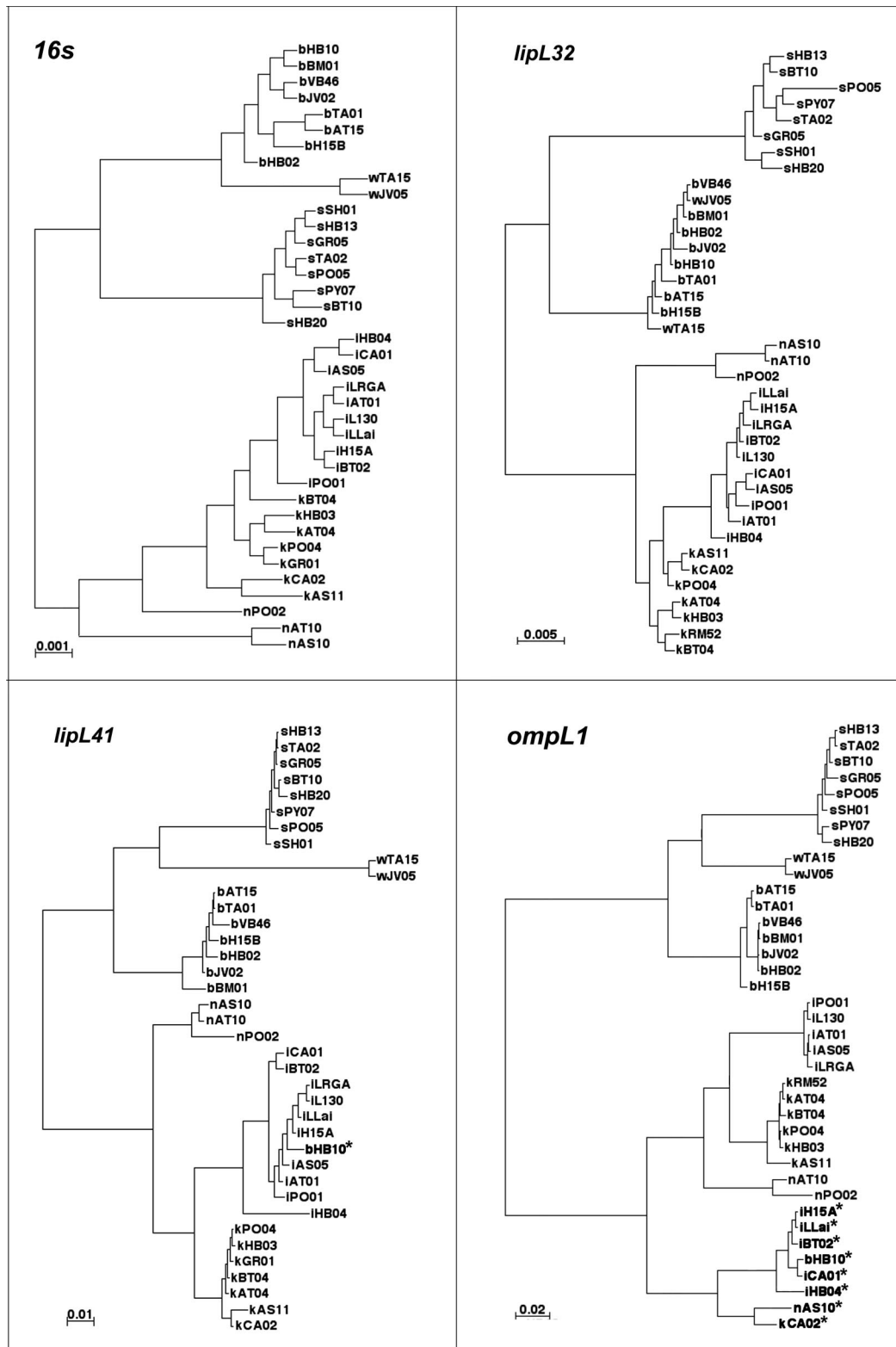
FIG. 3. Phylogenetic trees for the 16S, *lipL32*, *lipL41*, and *ompL1* genes. These unrooted phylogenetic trees summarize the posterior distribution of the evolutionary relationship among leptospiral strains inferred for the 16S, *lipL32*, *lipL41*, and *ompL1* genes with MrBayes. Sequences from the same 38 strains were used in each of the four trees. Each strain is represented in each gene tree by a lowercase letter (i, *L. interrogans*; k, *L. kirschneri*; n, *L. noguchii*; s, *L. santarosai*; b, *L. borgpetersenii*; w, *L. weilii*) followed by the 4-digit strain code (detailed descriptions of strains used in this study are given in Table 1). Sequences suggestive of horizontal transfer of DNA (bHB10* in the *lipL41* tree and the lowest branch of the *ompL1* tree) are indicated by asterisks. Numbers provided above major branches are branch lengths and are reported in the expected number of nucleotide substitutions per site.
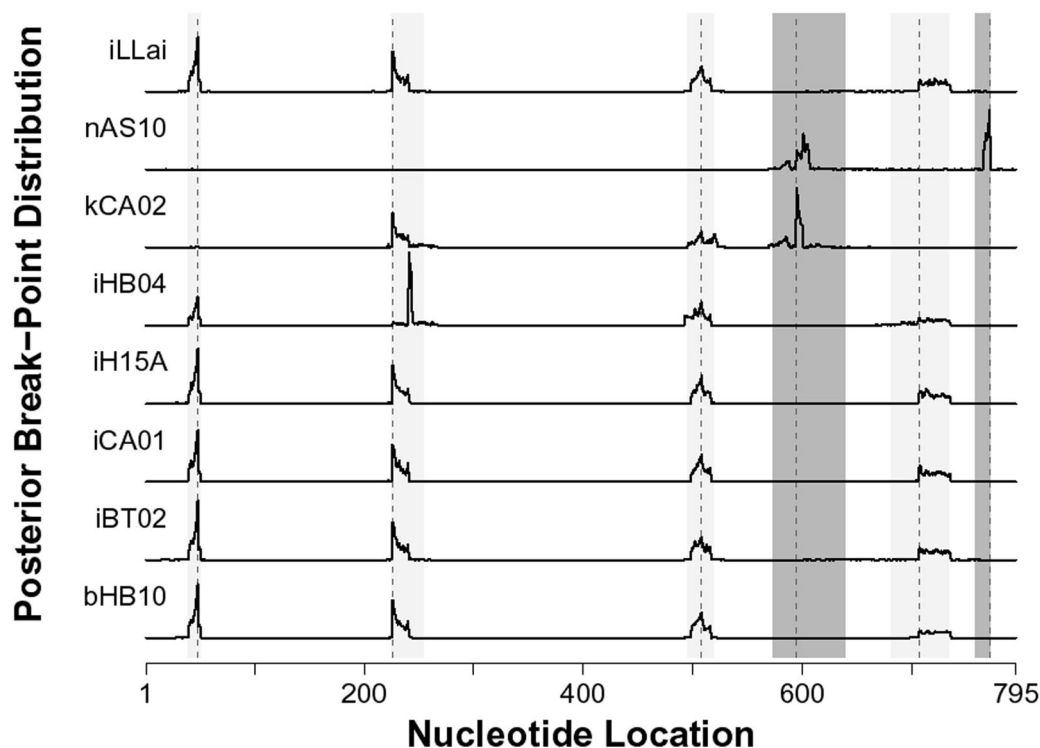
FIG. 4. Most probable breakpoint locations along the *ompL1* gene. The posterior distributions of breakpoint locations are plotted for 8 mosaic *ompL1* genes. Peaks on the distribution graphs indicate the most probable position of recombination. Three mosaic patterns are shown. The most common mosaic pattern has four breakpoints (shown by shaded bands covering their 95% Bayesian credible intervals). A second pattern, shown for nAS10, has two breakpoints (shown by dark shaded bands covering their 95% Bayesian credible intervals). The third pattern, occurring in kCA02, with three breakpoints, and is a hybrid of the first two patterns. Each strain is represented by a lowercase letter (i, *L. interrogans*; k, *L. kirschneri*; n, *L. noguchii*; b, *L. borgpetersenii*) followed by the 4-digit strain code (detailed descriptions of strains used in this study are given in Table 1).

the sequences (posterior probability, >0.99). This new branch is more closely related to the *L. interrogans*, *L. kirschneri*, and *L. noguchii* cluster than to the *L. santarosai*, *L. borgpetersenii*, and *L. weilii* cluster. Of the 8 *ompL1* sequences included in this study that are found in the new branch of the *ompL1* tree, five were from *L. interrogans* strains. The three other sequences found in the new branch of the *ompL1* tree were from members of the *L. kirschneri*, *L. noguchii*, and *L. borgpetersenii* species. Separate sequencing studies have also identified an *L. weilii* strain containing a similar *ompL1* gene sequence (data not shown), indicating the widespread distribution of DNA from the new branch of the *ompL1* tree.

Given the unexpected finding of a possibly new major lineage represented in only one of the gene trees, the *ompL1* sequence alignment was examined for an explanation. The individual site patterns from the different variable regions of the gene were compared. By eye, it was observed that sites from different regions tended to support different phylogenetic trees relating the 8 sequences in the new branch to the remaining major clades. Such spatial phylogenetic variation is suggestive of recombination (42). However, it remained unknown whether recombination between parentals derived from the typed clades resulted in progeny that were artifactually placed along a new major branch when assuming a single tree for the whole gene or whether recombination had occurred between

the typed clades and an allele derived from a new lineage. A BMCP model was used to formally test the hypothesis of recombination without assuming known parentals.

Multiple change point model. BMCP models were used to examine the evidence in favor of recombination, producing the sequences found in the new branch of the *ompL1* gene tree. To accomplish this task, 10 representative *ompL1* sequences with strong support of being nonrecombinants were identified. These representative sequences consisted of two sequences each from *L. interrogans*, *L. kirschneri*, *L. noguchii*, *L. santarosai*, and *L. borgpetersenii*. Chopping up the L (795) original alignment sites for the 10 sequences into M ($\approx$80) overlapping segments, each of length l (50) produced a list of P (50) unique trees. Assuming a prior probability of recombination of 0.5, it was found that, with >0.99 posterior probability, no recombination had occurred between the representative sequences, producing a single consistent tree along the entire gene alignment.

Using the 10 representative sequences as possible parentals, the BMCP model was then used to test whether the individual sequences found in the new branch of the *ompL1* tree were products of recombination. In all cases, the posterior probability of recombination was >0.99, revealing evidence of multiple recombination events producing the challenge sequences. Three basic mosaic patterns were identified. The most com-
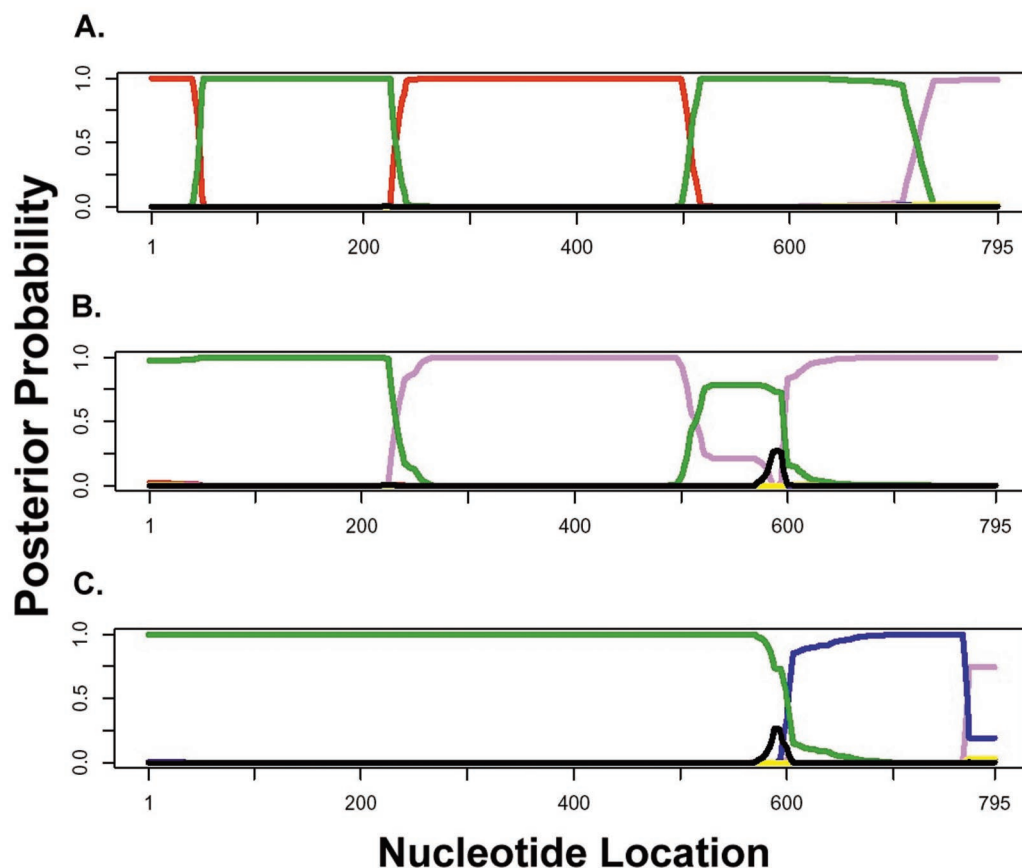
FIG. 5. Mosaic *ompL1* gene patterns. Three distinct mosaic *ompL1* gene patterns were revealed by the BMCP model. The posterior probabilities of various lineages (red, *L. interrogans*; purple, *L. kirschneri*; blue, *L. noguchii*; green, the peregrine allele) giving rise to a region are plotted for the *ompL1* genes from *L. interrogans* serovar Lai (A), *L. kirschneri* strain CA02 (B), and *L. noguchii* strain AS10 (C). The peregrine allele occurs in all three mosaic patterns between nucleotides 48 and 226 and between nucleotides 508 and 595.

mon pattern showed an *ompL1* gene with four breakpoints (most probable nucleotide locations, 48, 226, 508 and 707) (Fig. 4), dividing the sequence into five segments (Fig. 5A). The first and third segments were most closely related to *ompL1* sequences from *L. interrogans* strains. The fifth segment aligned with the *ompL1* sequences from *L. kirschneri*, and the second and fourth segments derived from a new lineage. The gene representing this new lineage will be referred to as a peregrine allele, referring to the ability of its DNA to migrate across species boundaries. The five-segment mosaic pattern was virtually identical in all five mosaic *ompL1* sequences found in *L. interrogans* strains and the mosaic *ompL1* sequences of *L. borgpetersenii* strain HB10 (Fig. 4).

The second pattern shows up only in the *ompL1* sequence found in *L. kirschneri* strain CA02 and contains three breakpoints, creating four segments (Fig. 5B). The first and second breakpoints in this second pattern occur at the same location as the second and third breakpoints sites in the five-segment *ompL1* mosaic (Fig. 4 and 5). However, the third breakpoint occurs in a new location near nucleotide 595. This four-segment *ompL1* mosaic contains segments derived from only two lineages: *L. kirschneri* and the peregrine allele. The first and third segments derived from the peregrine allele, and the sec-

ond and fourth segments were most closely related to *ompL1* sequences from *L. kirschneri*.

A third pattern was the *ompL1* sequence found in *L. noguchii* strain AS10. This was the simplest pattern, containing only two breakpoints and producing three segments (Fig. 5C). Each of the three segments was derived from a different lineage. The first segment was derived from the peregrine allele, the second segment retained the *L. noguchii* lineage, and the third segment aligned with *ompL1* sequences from *L. kirschneri*. The first breakpoint was the same as the third site in the four-segment *ompL1* mosaic of *L. kirschneri* strain CA02 (Fig. 4 and 5). However, the second breakpoint was unique to this mosaic pattern (most probable nucleotide location, 771).

Figure 5 shows that two *ompL1* regions contain DNA derived from the peregrine allele in each of the three mosaic patterns. The first segment derived from the peregrine allele is located between nucleotides 48 and 226 and corresponds to OmpL1 variable region 1 (Fig. 2A). The second segment derived from the peregrine allele is located between nucleotides 508 and 595 and corresponds to variable region 3 (Fig. 2A).

Not controlling for the effects of recombination can lead to a false-positive inference of positive selection acting on the codons within an alignment (1). Recombination was controlled

for when testing for positive selection by first partitioning the *ompL1* gene sequences into 5 separate segments. The most probable breakpoint locations for the first mosaic pattern were used to specify the borders between segments. For each segment, possible sites under positive selection were identified and the empirical Bayes probabilities were used to assess significance. Using this process, amino acid positions 260 and 262 (numbered from the start of the mature protein), both in the last segment, were identified as experiencing positive selection (posterior probabilities, >0.99 and 0.86, respectively). For comparison purposes, not controlling for recombination by assuming a single tree across all regions flagged one additional codon at amino acid position 81 (posterior probability, 0.84).

DISCUSSION

We describe multiple evolutionary mechanisms in leptospiral genes encoding OMPs, including nucleotide mutations producing single-nucleotide polymorphism and horizontal transfer of DNA between *Leptospira* species. Horizontal transfer of DNA appears to involve both entire genes (in the case of *lipL41*) and portions of genes (in the case of *ompL1*). Three different mosaic patterns of *ompL1* gene recombination were observed, varying from two to four sites of interspecies recombination involving DNA derived from two to three different leptospiral species origins per mosaic pattern. Other examples of mosaic genes encoding prokaryotic OMPs that have been found to undergo intragenic recombination include *porA* of *Neisseria meningitidis* (9), *porB* of *N. meningitidis* (2) and *Neisseria gonorrhoeae* (35), the intimin genes of *Escherichia coli* (30), and *ospC* (22) and *ospD* (27) of *Borrelia burgdorferi*. However, in those cases, no systematic approach to identification of the sites of recombination or the origin of recombined DNA was employed. The *ompL1* gene is the first prokaryotic gene in which these questions have been addressed with a BMCP model.

The eight strains containing mosaic *ompL1* genes belong to four different *Leptospira* species: *L. interrogans*, *L. kirschneri*, *L. noguchii*, and *L. borgpetersenii*. Similar mosaic *ompL1* genes have been found in an *L. weilii* strain (Fig. 5). The finding that all five of the mosaic *ompL1* sequences from *L. interrogans* strains were closely related, if not clonal, raises the possibility that the recombination events leading to the five-segment *ompL1* mosaic occurred (Fig. 4A) in a single ancestral *L. interrogans* strain. In addition, the same five-segment *ompL1* mosaic gene appears to have been horizontally transferred to *L. borgpetersenii* serovar Mini strain Sari (HB10), which is also the recipient of an intact (nonmosaicized) *L. interrogans lipL41* gene. The variation in breakpoint locations, segment sizes, and lineages in the three *ompL1* mosaic patterns indicates that recombination occurred independently in multiple ancestral strains. The *L. kirschneri* lineage of two of the four segments of the mosaic *ompL1* gene of *L. kirschneri* strain CA02 suggests that the recombination events leading to this four-segment mosaic occurred in this or an ancestral *L. kirschneri* strain. Likewise, the *L. noguchii* lineage of the second segment of the three-segment *ompL1* mosaic of *L. noguchii* strain AS10 suggests that the recombination events leading to this mosaic pattern occurred in an *L. noguchii* strain.

All eight *ompL1* mosaic genes and all three of the mosaic patterns observed contain DNA fragments that appear to have been derived from a novel *ompL1* allele, acquired by horizontal transfer of DNA. The source of this peregrine allele is uncertain. The mosaic sequences form an in-group in the *ompL1* phylogenetic tree (Fig. 3), suggesting a leptospiral pathogen as the source of the peregrine allele. One possibility is that the donor species is a previously undescribed member of the core group of pathogenic *Leptospira* species. An alternative explanation is that the source is a paralogous *ompL1* gene acquired from one of the known *Leptospira* species. The alternative explanation is less likely because Southern blot studies (13) and two leptospiral genome sequences reveal only a single *ompL1* locus per genome (32, 38).

In all three *ompL1* mosaic genes, the segments encoding the first and fourth surface-exposed loops are encoded by the peregrine allele. These portions of the *ompL1* gene appear to confer a selective advantage on leptospiral strains bearing mosaic *ompL1* genes. This is consistent with the finding of increased amino acid sequence variability in the *ompL1* gene regions corresponding to surface-exposed loops. The *ompL1* gene was examined for evidence of positive selection, and two amino acid sites in the fifth surface-exposed loop were identified (Fig. 2). The lack of more prevalent positive selection, particularly across additional variable loops, suggests that increased sequence variability in certain OmpL1 regions is not a result of immunological pressure. This is consistent with the idea that the renal tubules, the site of leptospiral colonization in the reservoir host, is an immunologically privileged location. The increased amino acid sequence variability may indicate that certain surface-exposed loop variants represent adaptations to the specific host environmental constraints or that amino acid changes are simply better tolerated in the surface-exposed loops.

It is unlikely that the intragenic recombination observed in some *ompL1* genes is artifactual. PCR products were directly sequenced without cloning, eliminating the possibility of point mutations introduced through amplification. Both strands of the PCR products were sequenced, reducing the possibility of sequencing errors. Although four different genes were sequenced in each of 38 strains, evidence for intragenic recombination was found only in *ompL1* genes. DNA recombination at the time of PCR amplification would be expected to result in a random set of recombination patterns. In contrast, the same pattern of intragenic recombination was found in six of the eight strains. One of the eight strains with evidence of *ompL1* intragenic recombination is the *L. interrogans* serovar Lai genome sequence strain, in which the *ompL1* sequence was derived from shotgun cloning of a clonal isolate. Finally, contamination of template DNA is also unlikely because the *Leptospira* donor species for some of the sequenced DNA has apparently never been isolated.

The studies presented here shed light on the phylogenetic organization of the genus *Leptospira*. The deeper branches of the *L. santarosai* and *L. weilii* monophyletic groups in most of the trees are consistent with the geographic isolation of these species to South America and Southeast Asia, respectively (5). In both the 16S and *lipL32* trees, the *L. interrogans* branch is a clonal branch emerging from within the *L. kirschneri* tree, suggesting that the *L. interrogans* species may have evolved from an ancestral *L. kirschneri* strain. The emergence of the *L.*

interrogans species relatively recently in leptospiral history would likely have been associated with changes in the ecology of its mammalian host. Since many *L. interrogans* strains are adapted to urban rodents, it is tempting to speculate that the expansion of the *L. interrogans* branch occurred with the development of agriculture.

Sequencing of the 16S gene is a standard approach for differentiating species in all branches of the phylogenetic tree of life (47). 16S sequences have been examined for the purpose of species differentiation within the genus *Leptospira* (19, 36) and have also been used as a method of determining the species identity of novel leptospiral strains (15). However, none of the genes examined here are ideal for this purpose when used individually. In the 16S tree, only *L. santarosai*, *L. borgpetersenii*, and *L. weilii* form monophyletic sister groups. In contrast, *L. interrogans* sequences form a monophyletic in-group within the *L. kirschneri* branch, and the *L. kirschneri* and *L. interrogans* sequences form a monophyletic group within the *L. noguchii* branch. Although the species of origin of most leptospiral 16S sequences could be readily identified by comparison with a large set of known sequences, certain sequences could be ambiguous. For example, the 16S sequence of *L. interrogans* strain PO01 would be difficult to distinguish from many *L. kirschneri* sequences. Due to greater sequence diversity, the *lipL32* phylogenetic tree provides greater differentiation of *L. noguchii*, *L. kirschneri*, *L. interrogans*, and *L. santarosai* sequences, but the *L. borgpetersenii* and *L. weilii* sequences are indistinguishable. All species form monophyletic sister groups in the *lipL41* and *ompL1* phylogenetic trees, but the potential for horizontal transfer of genes invalidates the use of these sequences as an independent means of species identification.

The DNA and amino acid sequences of the four genes examined in this study appear to be evolving at different rates. The 16S genes were the most conserved, followed by the *lipL32*, *lipL41*, and *ompL1* genes, listed in order of increasing sequence variability. The differences in sequence variability persisted even when only synonymous mutations were considered. Since LipL32 is the most highly expressed protein in pathogenic *Leptospira* species (14), codon bias is a possible explanation for the higher level of DNA sequence conservation. However, codon adaptation index measurements for *ompL1* and *lipL41* were higher than for *lipL32*, indicating that codon bias was not the explanation for the greater degree of sequence conservation among *lipL32* genes.

Shuttle vector, gene knockout, and complementation techniques have been developed for the saprophytic organism *Leptospira biflexa* (34). In contrast, strategies are not yet available for genetic manipulation of pathogenic *Leptospira* species. The studies presented here indicate that certain leptospiral strains are able to be transformed with foreign DNA and to undergo intragenic recombination. These studies also provide guidance for the development of OMPs as vaccines for leptospirosis. The immunological relevance of the increased amino acid sequence variation of OmpL1 can now be assessed. An OmpL1-based leptospiral vaccine might need to be polyvalent to account for the major OmpL1 variants. On the other hand, the high levels of sequence conservation indicate that monovalent LipL41- and/or LipL32-based vaccines have the potential for being broadly protective.

REFERENCES

1. **Anisimova, M., R. Nielsen, and Z. Yang.** 2003. Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. Genetics **164:**1229–1236.
2. **Bash, M. C., K. B. Lesiak, S. D. Banks, and C. E. Frasch.** 1995. Analysis of *Neisseria meningitidis* class 3 outer membrane protein gene variable regions and type identification using genetic techniques. Infect. Immun. **63:**1484–1490.
3. **Boursaux-Eude, C., I. Saint Girons, and R. Zuerner.** 1998. Leptospira genomics. Electrophoresis **19:**589–592.
4. **Branger, C., C. Sonrier, B. Chatrenet, B. Klonjkowski, N. Ruvoen-Clouet, A. Aubert, G. Andre-Fontaine, and M. Eloit.** 2001. Identification of the hemolysis-associated protein 1 as a cross-protective immunogen of *Leptospira interrogans* by adenovirus-mediated vaccination. Infect. Immun. **69:**6831–6838.
5. **Brenner, D. J., A. F. Kaufmann, K. R. Sulzer, A. G. Steigerwalt, F. C. Rogers, and R. S. Weyant.** 1999. Further determination of DNA relatedness between serogroups and serovars in the family *Leptospiraceae* with a proposal for *Leptospira alexanderi* sp. nov. and four new *Leptospira* genomospecies. Int. J. Syst. Bacteriol. **49:**839–858.
6. **Cullen, P. A., D. A. Haake, D. M. Bulach, R. L. Zuerner, and B. Adler.** 2003. LipL21 is a novel surface-exposed lipoprotein of pathogenic *Leptospira* species. Infect. Immun. **71:**2414–2421.
7. **de la Pena-Moctezuma, A., D. M. Bulach, T. Kalambaheti, and B. Adler.** 1999. Comparative analysis of the LPS biosynthetic loci of the genetic subtypes of serovar Hardjo: *Leptospira interrogans* subtype Hardjoprajitno and *Leptospira borgpetersenii* subtype Hardjobovis. FEMS Microbiol. Lett. **177:**319–326.
8. **Faine, S., B. Adler, C. Bolin, and P. Perolat.** 1999. *Leptospira* and leptospirosis, 2nd ed. MediSci, Melbourne, Australia.
9. **Feavers, I. M., A. B. Heath, J. A. Bygraves, and M. C. Maiden.** 1992. Role of horizontal genetic exchange in the antigenic variation of the class 1 outer membrane protein of *Neisseria meningitidis.* Mol. Microbiol. **6:**489–495.
10. **Flannery, B., D. Costa, F. P. Carvalho, H. Guerreiro, J. Matsunaga, E. D. Da Silva, A. G. Ferreira, L. W. Riley, M. G. Reis, D. A. Haake, and A. I. Ko.** 2001. Evaluation of recombinant *Leptospira* antigen-based enzyme-linked immunosorbent assays for the serodiagnosis of leptospirosis. J. Clin. Microbiol. **39:**3303–3310.
11. **Goldman, N., and Z. Yang.** 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol. **11:**725–736.
12. **Guerreiro, H., J. Croda, B. Flannery, M. Mazel, J. Matsunaga, M. Galvao Reis, P. N. Levett, A. I. Ko, and D. A. Haake.** 2001. Leptospiral proteins recognized during the humoral immune response to leptospirosis in humans. Infect. Immun. **69:**4958–4968.
13. **Haake, D. A., C. I. Champion, C. Martinich, E. S. Shang, D. R. Blanco, J. N. Miller, and M. A. Lovett.** 1993. Molecular cloning and sequence analysis of the gene encoding OmpL1, a transmembrane outer membrane protein of pathogenic *Leptospira* spp. J. Bacteriol. **175:**4225–4234.
14. **Haake, D. A., G. Chao, R. L. Zuerner, J. K. Barnett, D. Barnett, M. Mazel, J. Matsunaga, P. N. Levett, and C. A. Bolin.** 2000. The leptospiral major outer membrane protein LipL32 is a lipoprotein expressed during mammalian infection. Infect. Immun. **68:**2276–2285.
15. **Haake, D. A., M. Dundoo, R. Cader, B. M. Kubak, R. A. Hartskeerl, J. J. Sejvar, and D. A. Ashford.** 2002. Leptospirosis, water sports, and chemoprophylaxis. Clin. Infect. Dis. **34:**e40–43.
16. **Haake, D. A., C. Martinich, T. A. Summers, E. S. Shang, J. D. Pruetz, A. M. McCoy, M. K. Mazel, and C. A. Bolin.** 1998. Characterization of leptospiral outer membrane lipoprotein LipL36: downregulation associated with late-log-phase growth and mammalian infection. Infect. Immun. **66:**1579–1587.
17. **Haake, D. A., M. K. Mazel, A. M. McCoy, F. Milward, G. Chao, J. Matsunaga, and E. A. Wagar.** 1999. Leptospiral outer membrane proteins OmpL1 and LipL41 exhibit synergistic immunoprotection. Infect. Immun. **67:**6572–6582.
18. **Holder, M., and P. O. Lewis.** 2003. Phylogeny estimation: traditional and Bayesian approaches. Nat. Rev. Genet. **4:**275–284.
19. **Hookey, J. V., J. Bryden, and L. Gatehouse.** 1993. The use of 16S rDNA sequence analysis to investigate the phylogeny of *Leptospiraceae* and related spirochaetes. J. Gen. Microbiol. **139:**2585–2590.
20. **Huelsenbeck, J. P., and F. Ronquist.** 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:**754–755.
21. **Huelsenbeck, J. P., F. Ronquist, R. Nielsen, and J. P. Bollback.** 2001. Bayes-

ian inference of phylogeny and its impact on evolutionary biology. Science **294:**2310–2314.

22. **Jauris-Heipke, S., G. Liegl, V. Preac-Mursic, D. Rossler, E. Schwab, E. Soutschek, G. Will, and B. Wilske.** 1995. Molecular analysis of genes encoding outer surface protein C (OspC) of *Borrelia burgdorferi* sensu lato: relationship to *ospA* genotype and evidence of lateral gene exchange of ospC. J. Clin. Microbiol. **33:**1860–1866.

23. **Johnson, R. C., and V. G. Harris.** 1967. Differentiation of pathogenic and saprophytic letospires. I. Growth at low temperatures. J. Bacteriol. **94:**27–31.

24. **Kalambaheti, T., D. M. Bulach, K. Rajakumar, and B. Adler.** 1999. Genetic organization of the lipopolysaccharide O-antigen biosynthetic locus of *Leptospira borgpetersenii* serovar Hardjobovis. Microb. Pathog. **27:**105–117.

25. **Levett, P. N.** 2001. Leptospirosis. Clin. Microbiol. Rev. **14:**296–326.

26. **Li, W. H., M. Tanimura, and P. M. Sharp.** 1988. Rates and dates of divergence between AIDS virus nucleotide sequences. Mol. Biol. Evol. **5:**313–330.

27. **Marconi, R. T., D. S. Samuels, R. K. Landry, and C. F. Garon.** 1994. Analysis of the distribution and molecular heterogeneity of the *ospD* gene among the Lyme disease spirochetes: evidence for lateral gene exchange. J. Bacteriol. **176:**4572–4582.

28. **Matsunaga, J., M. A. Barocchi, J. Croda, T. A. Young, Y. Sanchez, I. Siqueira, C. A. Bolin, M. G. Reis, L. W. Riley, D. A. Haake, and A. I. Ko.** 2003. Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily. Mol. Microbiol. **49:**929–945.

29. **Matsunaga, J., T. A. Young, J. K. Barnett, D. Barnett, C. A. Bolin, and D. A. Haake.** 2002. Novel 45-kilodalton leptospiral protein that is processed to a 31-kilodalton growth-phase-regulated peripheral membrane protein. Infect. Immun. **70:**323–334.

30. **McGraw, E. A., J. Li, R. K. Selander, and T. S. Whittam.** 1999. Molecular evolution and mosaic structure of alpha, beta, and gamma intimins of pathogenic Escherichia coli. Mol. Biol. Evol. **16:**12–22.

31. **Nally, J. E., S. Artiushin, and J. F. Timoney.** 2001. Molecular characterization of thermoinduced immunogenic proteins Q1p42 and Hsp15 of *Leptospira interrogans.* Infect. Immun. **69:**7616–7624.

32. **Nascimento, A. L. T. O., A. I. Ko, E. A. L. Martins, C. B. Monteiro-Vitorello, P. L. Ho, D. A. Haake, S. Verjovski-Almeida, R. A. Hartskeerl, M. V. Marques, M. C. Oliveira, C. F. M. Menck, L. C. C. Leite, H. Carrer, L. L. Coutinho, W. M. Degrave, O. A. Dellagostin, H. El-Dorry, E. S. Ferro, M. I. T. Ferro, L. R. Furlan, M. Gamberin, E. A. Giglioti, A. Góes-Neto, G. H. Goldman, M. H. S. Goldman, R. Harakava, S. M. B. Jerônimo, I. L. M. Junqueira-de-Azevedo, E. T. Kimura, E. E. Kuramae, E. G. M. Lemos, M. V. F. Lemos, C. L. Marino, L. R. Nunes, R. C. de Oliveira, G. G. Pereira, M. S. Reis, A. Schriefer, W. J. Siqueira, P. Sommer, S. M. Tsai, A. J. G. Simpson, J. A. Ferro, L. E. A. Camargo, J. P. Kitajima, J. C. Setubal, and M. A. Van Sluys.** 2004. Comparative genomics of two Leptospira interrogans serovars reveals novel insights into physiology and pathogenesis. J. Bacteriol. **186:**2164–2172.

33. **Paraskevis, D., P. Lemey, M. Salemi, M. Suchard, Y. Van De Peer, and A. M. Vandamme.** 2003. Analysis of the evolutionary relationships of HIV-1 and SIVcpz sequences using Bayesian inference: implications for the origin of HIV-1. Mol. Biol. Evol. **20:**1986–1996.

34. **Picardeau, M., A. Brenot, and I. Saint Girons.** 2001. First evidence for gene replacement in *Leptospira* spp. inactivation of *L. biflexa flaB* results in nonmotile mutants deficient in endoflagella. Mol. Microbiol. **40:**189–199.

35. **Posada, D., K. A. Crandall, M. Nguyen, J. C. Demma, and R. P. Viscidi.** 2000. Population genetics of the *porB* gene of *Neisseria gonorrhoeae*: different dynamics in different homology groups. Mol. Biol. Evol. **17:**423–436.

36. **Postic, D., N. Riquelme-Sertour, F. Merien, P. Perolat, and G. Baranton.** 2000. Interest of partial 16S rDNA gene sequences to resolve heterogeneities between *Leptospira* collections: application to *L. meyeri.* Res. Microbiol. **151:**333–341.

37. **Ralph, D., and M. McClelland.** 1994. Phylogenetic evidence for horizontal transfer of an intervening sequence between species in a spirochete genus. J. Bacteriol. **176:**5982–5987.

38. **Ren, S. X., G. Fu, X. G. Jiang, R. Zeng, Y. G. Miao, H. Xu, Y. X. Zhang, H. Xiong, G. Lu, L. F. Lu, H. Q. Jiang, J. Jia, Y. F. Tu, J. X. Jiang, W. Y. Gu, Y. Q. Zhang, Z. Cai, H. H. Sheng, H. F. Yin, Y. Zhang, G. F. Zhu, M. Wan, H. L. Huang, Z. Qian, S. Y. Wang, W. Ma, Z. J. Yao, Y. Shen, B. Q. Qiang, Q. C. Xia, X. K. Guo, A. Danchin, I. Saint Girons, R. L. Somerville, Y. M. Wen, M. H. Shi, Z. Chen, J. G. Xu, and G. P. Zhao.** 2003. Unique physiological and pathogenic features of Leptospira interrogans revealed by whole-genome sequencing. Nature **422:**888–893.

39. **Shang, E. S., M. M. Exner, T. A. Summers, C. Martinich, C. I. Champion, R. E. Hancock, and D. A. Haake.** 1995. The rare outer membrane protein, OmpL1, of pathogenic *Leptospira* species is a heat-modifiable porin. Infect. Immun. **63:**3174–3181.

40. **Shang, E. S., T. A. Summers, and D. A. Haake.** 1996. Molecular cloning and sequence analysis of the gene encoding LipL41, a surface-exposed lipoprotein of pathogenic *Leptospira* species. Infect. Immun. **64:**2322–2330.

41. **Sinsheimer, J. S., M. A. Suchard, K. S. Dorman, F. Fang, and R. E. Weiss.** 2003. Are you my mother? Bayesian phylogenetic inference of recombination among putative parental strains. Appl. Bioinformatics **2:**131–144.

42. **Smith, J. M.** 1992. Analyzing the mosaic structure of genes. J. Mol. Evol. **34:**126–129.

43. **Suchard, M. A., R. E. Weiss, K. S. Dorman, and J. S. Sinsheimer.** 2003. Inferring spatial phylogenetic variation along nucleotide sequences: a multiple changepoint model. J. Am. Stat. Assoc. **98:**427–437.

44. **Suchard, M. A., R. E. Weiss, K. S. Dorman, and J. S. Sinsheimer.** 2002. Oh brother, where art thou? A Bayes factor test for recombination with uncertain heritage. Syst. Biol. **51:**715–728.

45. **Suchard, M. A., R. E. Weiss, and J. S. Sinsheimer.** 2001. Bayesian selection of continuous-time Markov chain evolutionary models. Mol. Biol. Evol. **18:**1001–1013.

46. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.

47. **Woese, C. R.** 2000. Interpreting the universal phylogenetic tree. Proc. Natl. Acad. Sci. USA **97:**8392–8396.

48. **Yang, Z.** 2002. Inference of selection from multiple species alignments. Curr. Opin. Genet. Dev. **12:**688–694.

49. **Yang, Z., and J. P. Bielawski.** 2000. Statistical methods for detecting molecular adaptation. Trends Ecol. Evol. **15:**496–503.

50. **Zuerner, R., D. Haake, B. Adler, and R. Segers.** 2000. Technological advances in the molecular biology of *Leptospira.* J. Mol. Microbiol. Biotechnol. **2:**455–462.

51. **Zuerner, R. L., and C. A. Bolin.** 1997. Differentiation of *Leptospira interrogans* isolates by IS1500 hybridization and PCR assays. J. Clin. Microbiol. **35:**2612–2617.

52. **Zuerner, R. L., J. L. Herrmann, and I. Saint Girons.** 1993. Comparison of genetic maps for two *Leptospira interrogans* serovars provides evidence for two chromosomes and intraspecies heterogeneity. J. Bacteriol. **175:**5445–5451.